

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Murphy, Brian R.
Collins, Peter L.
Whitehead, Stephen S.
Bukreyev, Alexander A.
Juhasz, Katalin
- (ii) TITLE OF INVENTION: PRODUCTION OF ATTENUATED RESPIRATORY
SYNCYTIAL VIRUS VACCINES FROM CLONED NUCLEOTIDE SEQUENCES
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Townsend and Townsend and Crew LLP
 - (B) STREET: Two Embarcadero Center, 8th Floor
 - (C) CITY: San Francisco
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94111-3834
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US
 - (B) FILING DATE: 15-JUL-1997
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/047,634
 - (B) FILING DATE: 23-MAY-1997
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/046,141
 - (B) FILING DATE: 09-MAY-1997
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/021,773
 - (B) FILING DATE: 15-JUL-1996
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Parmelee, Steven W.
 - (B) REGISTRATION NUMBER: 31,990
 - (C) REFERENCE/DOCKET NUMBER: 17634-000510
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 206-467-9600
 - (B) TELEFAX: 415-576-0300

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15223 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

| | |
|--|------|
| ACGCGAAAAA ATGCGTACAA CAAACTTGCA TAAACCAAAA AAATGGGGCA AATAAGAATT | 60 |
| TGATAAGTAC CACTTAAATT TAACTCCCTT GGTTAGAGAT GGGCAGCAAT TCATTGAGTA | 120 |
| TGATAAAAGT TAGATTACAA AATTGTGTTG ACAATGATGA AGTAGCATTG TTAAAAATAA | 180 |
| CATGCTATAC TGATAAATTA ATACATTTAA CTAATGCTTT GGCTAAGGCA GTGATACATA | 240 |
| CAATCAAATT GAATGGCATT GTGTTTGTGC ATGTTATTAC AAGTAGTGAT ATTTGCCCTA | 300 |
| ATAATAATAT TGTAATAAAA TCCAATTTCA CAACAATGCC AGTACTACAA AATGGAGGTT | 360 |
| ATATATGGGA AATGATGGAA TTAACACATT GCTCTCAACC TAATGGTCTA CTAGATGACA | 420 |
| ATTGTGAAAT TAAATTCTCC AAAAACTAA GTGATTCAAC AATGACCAAT TATATGAATC | 480 |
| AATTATCTGA ATTACTTGGA TTTGATCTTA ATCCATAAAT TATAATTAAT ATCAACTAGC | 540 |
| AAATCAATGT CACTAACACC ATTAGTTAAT ATAAAACTTA ACAGAAGACA AAAATGGGGC | 600 |
| AAATAAATCA ATTCAGCCAA CCCAACCATG GACACAACCC ACAATGATAA TACACCACAA | 660 |
| AGACTGATGA TCACAGACAT GAGACCGTTG TCACTTGAGA CCATAATAAC ATCACTAACC | 720 |
| AGAGACATCA TAACACACAA ATTTATATAC TTGATAAATC ATGAATGCAT AGTGAGAAAA | 780 |
| CTTGATGAAA AGCAGGCCAC ATTTACATTC CTGGTCAACT ATGAAATGAA ACTATTACAC | 840 |
| AAAGTAGGAA GCACTAAATA TAAAAAATAT ACTGAATACA ACACAAAATA TGGCACTTTC | 900 |
| CCTATGCCAA TATTCATCAA TCATGATGGG TTCTTAGAAT GCATTGGCAT TAAGCCTACA | 960 |
| AAGCATACTC CCATAATATA CAAGTATGAT CTCAATCCAT AAATTTCAAC ACAATATTCA | 1020 |
| CACAATCTAA AACAACAACCT CTATGCATAA CTATACTCCA TAGTCCAGAT GGAGCCTGAA | 1080 |
| AATTATAGTA ATTTAAACT TAAGGAGAGA TATAAGATAG AAGATGGGGC AAATACAACC | 1140 |
| ATGGCTCTTA GCAAAGTCAA GTTGAATGAT AACTCAACA AAGATCAACT TCTGTCATCC | 1200 |
| AGCAAATACA CCATCCAACG GAGCACAGGA GATAGTATTG ATACTCCTAA TTATGATGTG | 1260 |
| CAGAAACACA TCAATAAGTT ATGTGGCATG TTATTAATCA CAGAAGATGC TAATCATAAA | 1320 |
| TTCACTGGGT TAATAGGTAT GTTATATGCG ATGTCTAGGT TAGGAAGAGA AGACACCATA | 1380 |
| AAAATACTCA GAGATGCGGG ATATCATGTA AAAGCAAATG GAGTAGATGT AACACACAT | 1440 |
| CGTCAAGACA TTAATGGAAA AGAAATGAAA TTTGAAGTGT TAACATTGGC AAGCTTAACA | 1500 |
| ACTGAAATTC AAATCAACAT TGAGATAGAA TCTAGAAAAT CCTACAAAAA AATGCTAAAA | 1560 |
| GAAATGGGAG AGGTAGCTCC AGAATACAGG CATGACTCTC CTGATTGTGG GATGATAATA | 1620 |
| TTATGTATAG CAGCATTAGT AATAACTAAA TTAGCAGCAG GGGACAGATC TGGTCTTACA | 1680 |
| GCCGTGATTA GGAGAGCTAA TAATGTCCTA AAAAATGAAA TGAAACGTTA CAAAGGCTTA | 1740 |
| CTACCCAAGG ACATAGCCAA CAGCTTCTAT GAAGTGTGTTG AAAAACATCC CCACTTTATA | 1800 |
| GATGTTTTTG TTCATTTTGG TATAGCACAA TCTTCTACCA GAGGTGGCAG TAGAGTTGAA | 1860 |
| GGGATTTTTG CAGGATTGTT TATGAATGCC TATGGTGCAG GGCAAGTGAT GTTACGGTGG | 1920 |
| GGAGTCTTAG CAAAATCAGT TAAAAATATT ATGTTAGGAC ATGCTAGTGT GCAAGCAGAA | 1980 |
| ATGGAACAAG TTGTTGAGGT TTATGAATAT GCCCAAAAAT TGGGTGGTGA AGCAGGATTC | 2040 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| TACCATATAT | TGAACAACCC | AAAAGCATCA | TTATTATCTT | TGACTCAATT | TCCTCACTTC | 2100 |
| TCCAGTGTAG | TATTAGGCAA | TGCTGCTGGC | CTAGGCATAA | TGGGAGAGTA | CAGAGGTACA | 2160 |
| CCGAGGAATC | AAGATCTATA | TGATGCAGCA | AAGGCATATG | CTGAACAAC | CAAAGAAAAT | 2220 |
| GGTGTGATTA | ACTACAGTGT | ACTAGACTTG | ACAGCAGAAG | AACTAGAGGC | TATCAAACAT | 2280 |
| CAGCTTAATC | CAAAAGATAA | TGATGTAGAG | CTTTGAGTTA | ATAAAAAATG | GGGCAAATAA | 2340 |
| ATCATCATGG | AAAAGTTTGC | TCCTGAATTC | CATGGAGAAG | ATGCAAACAA | CAGGGCTACT | 2400 |
| AAATTCCTAG | AATCAATAAA | GGGCAAATTC | ACATCACCCA | AAGATCCCAA | GAAAAAAGAT | 2460 |
| AGTATCATAT | CTGTCAACTC | AATAGATATA | GAAGTAACCA | AAGAAAGCCC | TATAACATCA | 2520 |
| AATTCAACTA | TTATCAACCC | AACAAATGAG | ACAGATGATA | CTGCAGGGAA | CAAGCCCAAT | 2580 |
| TATCAAAGAA | AACCTCTAGT | AAGTTTCAAA | GAAGACCCTA | CACCAAGTGA | TAATCCCTTT | 2640 |
| TCTAAACTAT | ACAAAGAAAC | CATAGAAACA | TTTGATAACA | ATGAAGAAGA | ATCCAGCTAT | 2700 |
| TCATACGAAG | AAATAAATGA | TCAGACAAAC | GATAATATAA | CAGCAAGATT | AGATAGGATT | 2760 |
| GATGAAAAAT | TAAGTGAAAT | ACTAGGAATG | CTTCACACAT | TAGTAGTGGC | AAGTGCAGGA | 2820 |
| CCTACATCTG | CTCGGGATGG | TATAAGAGAT | GCCATGGTTG | GTTTAAGAGA | AGAAATGATA | 2880 |
| GAAAAAATCA | GAAGTGAAGC | ATTAATGACC | AATGACAGAT | TAGAAGCTAT | GGCAAGACTC | 2940 |
| AGGAATGAGG | AAAGTGAAAA | GATGGCAAAA | GACACATCAG | ATGAAGTGTC | TCTCAATCCA | 3000 |
| ACATCAGAGA | AATTGAACAA | CCTATTGGAA | GGGAATGATA | GTGACAATGA | TCTATCACTT | 3060 |
| GAAGATTTCT | GATTAGTTAC | CAATCTTCAC | ATCAACACAC | AATACCAACA | GAAGACCAAC | 3120 |
| AAACTAACCA | ACCCAATCAT | CCAACCAAAC | ATCCATCCGC | CAATCAGCCA | AACAGCCAAC | 3180 |
| AAAACAACCA | GCCAATCCAA | AACTAACCAC | CCGGAAAAAA | TCTATAATAT | AGTTACAAAA | 3240 |
| AAAGGAAAGG | GTGGGGCAAA | TATGGAAACA | TACGTGAACA | AGCTTCACGA | AGGCTCCACA | 3300 |
| TACACAGCTG | CTGTTCAATA | CAATGTCTTA | GAAAAAGACG | ATGACCCTGC | ATCACTTACA | 3360 |
| ATATGGGTGC | CCATGTTCCA | ATCATCTATG | CCAGCAGATT | TACTTATAAA | AGAACTAGCT | 3420 |
| AATGTCAACA | TACTAGTGAA | ACAAATATCC | ACACCCAAGG | GACCTTCACT | AAGAGTCATG | 3480 |
| ATAAACTCAA | GAAGTGCAGT | GCTAGCACAA | ATGCCCAGCA | AATTTACCAT | ATGCGCTAAT | 3540 |
| GTGTCCTTGG | ATGAAAGAAG | CAAACTAGCA | TATGATGTAA | CCACACCCTG | TGAAATCAAG | 3600 |
| GCATGTAGTC | TAACATGCCT | AAAATCAAAA | AATATGTTGA | CTACAGTTAA | AGATCTCACT | 3660 |
| ATGAAGACAC | TCAACCCTAC | ACATGATATT | ATTGCTTTAT | GTGAATTTGA | AAACATAGTA | 3720 |
| ACATCAAAAA | AAGTCATAAT | ACCAACATAC | CTAAGATCCA | TCAGTGTGAG | AAATAAGAT | 3780 |
| CTGAACACAC | TTGAAAATAT | AACAACCACT | GAATTCAAAA | ATGCTATCAC | AAATGCAAAA | 3840 |
| ATCATCCCTT | ACTCAGGATT | ACTATTAGTC | ATCACAGTGA | CTGACAACAA | AGGAGCATTC | 3900 |
| AAATACATAA | AGCCACAAAG | TCAATTCATA | GTAGATCTTG | GAGCTTACCT | AGAAAAAGAA | 3960 |
| AGTATATATT | ATGTTACCAC | AAATTGGAAG | CACACAGCTA | CACGATTTGC | AATCAAACCC | 4020 |
| ATGGAAGATT | AACCTTTTTT | CTCTACATCA | GTGTGTTAAT | TCATACAAAC | TTTCTACCTA | 4080 |
| CATTCTTCAC | TTCACCATCA | CAATCACAAA | CACTCTGTGG | TTCAACCAAT | CAAACAAAAC | 4140 |

| | | | | | | |
|------------|------------|------------|------------|-------------|-------------|------|
| TTATCTGAAG | TCCCAGATCA | TCCCAAGTCA | TTGTTTATCA | GATCTAGTAC | TCAAATAAGT | 4200 |
| TAATAAAAAA | TATACACATG | GGGCAAATAA | TCATTGGAGG | AAATCCAAC | AATCACAATA | 4260 |
| TCTGTTAACA | TAGACAAGTC | CACACACCAT | ACAGAATCAA | CCAATGGAAA | ATACATCCAT | 4320 |
| AACAATAGAA | TTCTCAAGCA | AATTCTGGCC | TTACTTTACA | CTAATACACA | TGATCACAAC | 4380 |
| AATAATCTCT | TTGCTAATCA | TAATCTCCAT | CATGATTGCA | ATACTAAACA | AACCTTTGTGA | 4440 |
| ATATAACGTA | TTCCATAACA | AAACCTTTGA | GTTACCAAGA | GCTCGAGTCA | ACACATAGCA | 4500 |
| TTCATCAATC | CAACAGCCCA | AAACAGTAAC | CTTGCATTTA | AAAATGAACA | ACCCCTACCT | 4560 |
| CTTTACAACA | CCTCATTAAC | ATCCCACCAT | GCAAACCACT | ATCCATACTA | TAAAGTAGTT | 4620 |
| AATTAAAAAT | AGTCATAACA | ATGAACTAGG | ATATCAAGAC | TAACAATAAC | ATTGGGGCAA | 4680 |
| ATGCAAACAT | GTCCAAAAAC | AAGGACCAAC | GCACCGCTAA | GACATTAGAA | AGGACCTGGG | 4740 |
| ACACTCTCAA | TCATTTATTA | TTCATATCAT | CGTGCTTATA | TAAGTTAAAT | CTTAAATCTG | 4800 |
| TAGCACAAAT | CACATTATCC | ATTCTGGCAA | TGATAATCTC | AACCTTCACTT | ATAATTGCAG | 4860 |
| CCATCATATT | CATAGCCTCG | GCAAACCACA | AAGTCACACC | AACAAC | ATCATAAAG | 4920 |
| ATGCAACAAG | CCAGATCAAG | AACACAACCC | CAACATACCT | CACCCAGAAT | CCTCAGCTTG | 4980 |
| GAATCAGTCC | CTCTAATCCG | TCTGAAATTA | CATCACAAAT | CACCACCATA | CTAGCTTCAA | 5040 |
| CAACACCAGG | AGTCAAGTCA | ACCCTGCAAT | CCACAACAGT | CAAGACCAAA | AACACAACAA | 5100 |
| CAACTCAAAC | ACAACCCAGC | AAGCCCACCA | CAAAACAACG | CCAAAACAAA | CCACCAAGCA | 5160 |
| AACCCAATAA | TGATTTTCAC | TTTGAAGTGT | TCAACTTTGT | ACCCTGCAGC | ATATGCAGCA | 5220 |
| ACAATCCAAC | CTGCTGGGCT | ATCTGCAAAA | GAATACCAAA | CAAAAAACCA | GGAAAGAAAA | 5280 |
| CCACTACCAA | GCCCACAAAA | AAACCAACCC | TCAAGACAAC | CAAAAAAGAT | CCCAAACCTC | 5340 |
| AAACCACTAA | ATCAAAGGAA | GTACCCACCA | CCAAGCCCAC | AGAAGAGCCA | ACCATCAACA | 5400 |
| CCACCAAAAC | AAACATCATA | ACTACACTAC | TCACCTCCAA | CACCACAGGA | AATCCAGAAC | 5460 |
| TCACAAGTCA | AATGGAAACC | TTCCACTCAA | CTTCCTCCGA | AGGCAATCCA | AGCCCTTCTC | 5520 |
| AAGTCTCTAC | AACATCCGAG | TACCCATCAC | AACCTTCATC | TCCACCCAAC | ACACCACGCC | 5580 |
| AGTAGTTACT | TAAAAACATA | TTATCACAAA | AGGCCTTGAC | CAACTTAAAC | AGAATCAAAA | 5640 |
| TAAACTCTGG | GGCAAATAAC | AATGGAGTTG | CTAATCCTCA | AAGCAAATGC | AATTACCACA | 5700 |
| ATCCTCACTG | CAGTCACATT | TTGTTTTGCT | TCTGGTCAAA | ACATCACTGA | AGAATTTTAT | 5760 |
| CAATCAACAT | GCAGTGCAGT | TAGCAAAGGC | TATCTTAGTG | CTCTGAGAAC | TGGTTGGTAT | 5820 |
| ACCAGTGTTA | TAATATAGA | ATTAAGTAAT | ATCAAGAAAA | ATAAGTGTA | TGGAACAGAT | 5880 |
| GCTAAGGTAA | AATTGATAAA | ACAAGAATTA | GATAAATATA | AAAATGCTGT | AACAGAATTG | 5940 |
| CAGTTGCTCA | TGCAAAGCAC | ACAAGCAACA | AACAATCGAG | CCAGAAGAGA | ACTACCAAGG | 6000 |
| TTTATGAATT | ATACACTCAA | CAATGCCAAA | AAAACCAATG | TAACATTAAG | CAAGAAAAGG | 6060 |
| AAAAGAAGAT | TTCTTGGTTT | TTTGTTAGGT | GTTGGATCTG | CAATCGCCAG | TGGCGTTGCT | 6120 |
| GTATCTAAGG | TCCTGCACCT | AGAAGGGGAA | GTGAACAAGA | TCAAAAGTGC | TCTACTATCC | 6180 |
| ACAAACAAGG | CTGTAGTCAG | CTTATCAAAT | GGAGTTAGTG | TTTTAACCAG | CAAAGTGTTA | 6240 |

| | | | | | | |
|------------|-------------|------------|-------------|------------|------------|------|
| GACCTCAAAA | ACTATATAGA | TAAACAATTG | TTACCTATTG | TGAACAAGCA | AAGCTGCAGC | 6300 |
| ATATCAAATA | TAGAACTGT | GATAGAGTTC | CAACAAAAGA | ACAACAGACT | ACTAGAGATT | 6360 |
| ACCAGGGAAT | TTAGTGTTAA | TGCAGGCGTA | ACTACACCTG | TAAGCACTTA | CATGTTAACT | 6420 |
| AATAGTGAAT | TATTGTCATT | AATCAATGAT | ATGCCTATAA | CAAATGATCA | GAAAAAGTTA | 6480 |
| ATGTCCAACA | ATGTTCAAAT | AGTTAGACAG | CAAAGTTACT | CTATCATGTC | CATAATAAAA | 6540 |
| GAGGAAGTCT | TAGCATATGT | AGTACAATTA | CCACTATATG | GTGTTATAGA | TACACCCTGT | 6600 |
| TGGAAACTAC | ACACATCCCC | TCTATGTACA | ACCAACACAA | AAGAAGGGTC | CAACATCTGT | 6660 |
| TTAACAAGAA | CTGACAGAGG | ATGGTACTGT | GACAATGCAG | GATCAGTATC | TTTCTTCCCA | 6720 |
| CAAGCTGAAA | CATGTAAAGT | TCAATCAAAT | CGAGTATTTT | GTGACACAAT | GAACAGTTTA | 6780 |
| ACATTACCAA | GTGAAGTAAA | TCTCTGCAAT | GTTGACATAT | TCAACCCCAA | ATATGATTGT | 6840 |
| AAAATTATGA | CTTCAAAAAC | AGATGTAAGC | AGCTCCGTTA | TCACATCTCT | AGGAGCCATT | 6900 |
| GTGTCATGCT | ATGGCAAAAAC | TAAATGTACA | GCATCCAATA | AAAATCGTGG | AATCATAAAG | 6960 |
| ACATTTTCTA | ACGGGTGCGA | TTATGTATCA | AATAAAGGGG | TGGACACTGT | GTCTGTAGGT | 7020 |
| AACACATTAT | ATTATGTAAA | TAAGCAAGAA | GGTAAAAGTC | TCTATGTAAA | AGGTGAACCA | 7080 |
| ATAATAAATT | TCTATGACCC | ATTAGTATTC | CCCTCTGATG | AATTTGATGC | ATCAATATCT | 7140 |
| CAAGTCAACG | AGAAGATTAA | CCAGAGCCTA | GCATTTATTC | GTAAATCCGA | TGAATTATTA | 7200 |
| CATAATGTAA | ATGCTGGTAA | ATCCACCACA | AATATCATGA | TAACTACTAT | AATTATAGTG | 7260 |
| ATTATAGTAA | TATTGTTATC | ATTAATTGCT | GTTGGACTGC | TCTTATACTG | TAAGGCCAGA | 7320 |
| AGCACACCAG | TCACACTAAG | CAAAGATCAA | CTGAGTGGTA | TAAATAATAT | TGCATTTAGT | 7380 |
| AACTAAATAA | AAATAGCACC | TAATCATGTT | CTTACAATGG | TTTACTATCT | GCTCATAGAC | 7440 |
| AACCCATCTG | TCATTGGATT | TTCTTAAAT | CTGAAC TTCA | TCGAAACTCT | CATCTATAAA | 7500 |
| CCATCTCACT | TACACTATTT | AAGTAGATTC | CTAGTTTATA | GTTATATAAA | ACACAATTGC | 7560 |
| ATGCCAGATT | AACTTACCAT | CTGTAAAAAT | GAAAAC TGGG | GCAAATATGT | CACGAAGGAA | 7620 |
| TCCTTGCAAA | TTTGAAATTC | GAGGTCATTG | CTTAAATGGT | AAGAGGTGTC | ATTTTAGTCA | 7680 |
| TAATTATTTT | GAATGGCCAC | CCCATGCACT | GCTTGTAAGA | CAAAACTTTA | TGTTAAACAG | 7740 |
| AATACTTAAG | TCTATGGATA | AAAGTATAGA | TACCTTATCA | GAAATAAGTG | GAGCTGCAGA | 7800 |
| GTTGGACAGA | ACAGAAGAGT | ATGCTCTTGG | TGTAGTTGGA | GTGCTAGAGA | GTTATATAGG | 7860 |
| ATCAATAAAC | AATATAACTA | AACAATCAGC | ATGTGTTGCC | ATGAGCAAAC | TCCTCACTGA | 7920 |
| ACTCAATAGT | GATGATATCA | AAAAGCTGAG | GGACAATGAA | GAGCTAAATT | CACCCAAGAT | 7980 |
| AAGAGTGTAC | AATACTGTCA | TATCATATAT | TGAAAGCAAC | AGGAAAAACA | ATAAACAAAC | 8040 |
| TATCCATCTG | TTAAAAAGAT | TGCCAGCAGA | CGTATTGAAG | AAAACCATCA | AAAACACATT | 8100 |
| GGATATCCAT | AAGAGCATAA | CCATCAACAA | CCCAAAGAA | TCAACTGTTA | GTGATACAAA | 8160 |
| TGACCATGCC | AAAAATAATG | ATACTACCTG | ACAAATATCC | TTGTAGTATA | ACTTCCATAC | 8220 |
| TAATAACAAG | TAGATGTAGA | GTTACTATGT | ATAATCAAAA | GAACACACTA | TATTTCAATC | 8280 |
| AAAACAACCC | AAATAACCAT | ATGTACTCAC | CGAATCAAAC | ATTCAATGAA | ATCCATTGGA | 8340 |

| | | | | | | |
|-------------|------------|-------------|------------|------------|-------------|-------|
| CCTCTCAAGA | ATTGATTGAC | ACAATTCAAA | ATTTTCTACA | ACATCTAGGT | ATTATTGAGG | 8400 |
| ATATATATAC | AATATATATA | TTAGTGTCAT | AACACTCAAT | TCTAACACTC | ACCACATCGT | 8460 |
| TACATTATTA | ATTCAAACAA | TTCAAGTTGT | GGGACAAAAT | GGATCCCATT | ATTAATGGAA | 8520 |
| ATTCTGCTAA | TGTTTATCTA | ACCGATAGTT | ATTTAAAAGG | TGTTATCTCT | TTCTCAGAGT | 8580 |
| GTAATGCTTT | AGGAAGTTAC | ATATTCAATG | GTCCTTATCT | CAAAAATGAT | TATACCAACT | 8640 |
| TAATTAGTAG | ACAAAATCCA | TTAATAGAAC | ACATGAATCT | AAAGAACTA | AATATAACAC | 8700 |
| AGTCCTTAAT | ATCTAAGTAT | CATAAAGGTG | AAATAAAATT | AGAAGAACCT | ACTTATTTTC | 8760 |
| AGTCATTACT | TATGACATAC | AAGAGTATGA | CCTCGTCAGA | ACAGATTGCT | ACCACTAATT | 8820 |
| TACTTAAAAA | GATAATAAGA | AGAGCTATAG | AAATAAGTGA | TGTCAAAGTC | TATGCTATAT | 8880 |
| TGAATAAACT | AGGGCTTAAA | GAAAAGGACA | AGATTAAATC | CAACAATGGA | CAAGATGAAG | 8940 |
| ACAACTCAGT | TATTACGACC | ATAATCAAAG | ATGATATACT | TTCAGCTGTT | AAAGATAATC | 9000 |
| AATCTCATCT | TAAAGCAGAC | AAAAATCACT | CTACAAAACA | AAAAGACACA | ATCAAAAACAA | 9060 |
| CACTCTTGAA | GAAATTGATG | TGTTCAATGC | AACATCCTCC | ATCATGGTTA | ATACATTGGT | 9120 |
| TTAACTTATA | CACAAAATTA | AACAACATAT | TAACACAGTA | TCGATCAAAT | GAGGTAAAAA | 9180 |
| ACCATGGGTT | TACATTGATA | GATAATCAAA | CTCTTAGTGG | ATTTCAATTT | ATTTTGAACC | 9240 |
| AATATGGTTG | TATAGTTTAT | CATAAGGAAC | TCAAAAGAAT | TACTGTGACA | ACCTATAATC | 9300 |
| AATTCTTGAC | ATGGAAAGAT | ATTAGCCTTA | GTAGATTAAA | TGTTTGTTTA | ATTACATGGA | 9360 |
| TTAGTAACTG | CTTGAACACA | TTAAATAAAA | GCTTAGGCTT | AAGATGCGGA | TTCAATAATG | 9420 |
| TTATCTTGAC | ACAACTATTC | CTTTATGGAG | ATTGTATACT | AAAGCTATTT | CACAATGAGG | 9480 |
| GGTTCTACAT | AATAAAAGAG | GTAGAGGGAT | TTATTATGTC | TCTAATTTTA | AATATAACAG | 9540 |
| AAGAAGATCA | ATTCAGAAAA | CGATTTTATA | ATAGTATGCT | CAACAACATC | ACAGATGCTG | 9600 |
| CTAATAAAGC | TCAGAAAAAT | CTGCTATCAA | GAGTATGTCA | TACATTATTA | GATAAGACAG | 9660 |
| TGTCCGATAA | TATAATAAAT | GGCAGATGGA | TAATTCTATT | AAGTAAGTTC | CTTAAATTAA | 9720 |
| TTAAGCTTGC | AGGTGACAAT | AACCTTAACA | ATCTGAGTGA | ACTATATTTT | TTGTTTCAGAA | 9780 |
| TATTTGGACA | CCCAATGGTA | GATGAAAGAC | AAGCCATGGA | TGCTGTTAAA | ATTAATTGCA | 9840 |
| ATGAGACCAA | ATTTTACTTG | TTAAGCAGTC | TGAGTATGTT | AAGAGGTGCC | TTTATATATA | 9900 |
| GAATTATAAA | AGGGTTTGTA | AATAATTACA | ACAGATGGCC | TACTTTAAGA | AATGCTATTG | 9960 |
| TTTTACCCTT | AAGATGGTTA | ACTTACTATA | AACTAAACAC | TTATCCTTCT | TTGTTGGAAC | 10020 |
| TTACAGAAAG | AGATTTGATT | GTGTTATCAG | GACTACGTTT | CTATCGTGAG | TTTCGGTTGC | 10080 |
| CTAAAAAAGT | GGATCTTGAA | ATGATTATAA | ATGATAAAGC | TATATCACCT | CCTAAAAATT | 10140 |
| TGATATGGAC | TAGTTTCCCT | AGAAATTACA | TGCCATCACA | CATACAAAAC | TATATAGAAC | 10200 |
| ATGAAAAATT | AAAATTTTCC | GAGAGTGATA | AATCAAGAAG | AGTATTAGAG | TATTATTTAA | 10260 |
| GAGATAACAA | ATTCAATGAA | TGTGATTTAT | ACAACTGTGT | AGTTAATCAA | AGTTATCTCA | 10320 |
| ACAACCECTAA | TCATGTGGTA | TCATTGACAG | GCAAAGAAAG | AGAACTCAGT | GTAGGTAGAA | 10380 |
| TGTTTGCAAT | GCAACCGGGA | ATGTTTCAGAC | AGGTTCAAAT | ATTGGCAGAG | AAAATGATAG | 10440 |

| | | | | | | |
|------------|------------|------------|------------|-------------|-------------|-------|
| CTGAAAACAT | TTTACAATTC | TTTCCTGAAA | GTCTTACAAG | ATATGGTGAT | CTAGAACTAC | 10500 |
| AAAAAATATT | AGAACTGAAA | GCAGGAATAA | GTAACAAATC | AAATCGCTAC | AATGATAATT | 10560 |
| ACAACAATTA | CATTAGTAAG | TGCTCTATCA | TCACAGATCT | CAGCAAATTC | AATCAAGCAT | 10620 |
| TTCGATATGA | AACGTCATGT | ATTTGTAGTG | ATGTGCTGGA | TGAACTGCAT | GGTGTAACAAT | 10680 |
| CTCTATTTTC | CTGGTTACAT | TTAACTATTC | CTCATGTCAC | AATAATATGC | ACATATAGGC | 10740 |
| ATGCACCCCC | CTATATAGGA | GATCATATTG | TAGATCTTAA | CAATGTAGAT | GAACAAAGTG | 10800 |
| GATTATATAG | ATATCACATG | GGTGGCATCG | AAGGGTGGTG | TCAAAAACCTA | TGGACCATAG | 10860 |
| AAGCTATATC | ACTATTGGAT | CTAATATCTC | TCAAAGGGAA | ATTCTCAATT | ACTGCTTTAA | 10920 |
| TTAATGGTGA | CAATCAATCA | ATAGATATAA | GCAAACCAAT | CAGACTCATG | GAAGGTCAAA | 10980 |
| CTCATGCTCA | AGCAGATTAT | TTGCTAGCAT | TAAATAGCCT | TAAATTACTG | TATAAAGAGT | 11040 |
| ATGCAGGCAT | AGGCCACAAA | TTAAAAGGAA | CTGAGACTTA | TATATCACGA | GATATGCAAT | 11100 |
| TTATGAGTAA | AACAATTCAA | CATAACGGTG | TATATTACCC | AGCTAGTATA | AAGAAAGTCC | 11160 |
| TAAGAGTGGG | ACCGTGGATA | AACACTATAC | TTGATGATTT | CAAAGTGAGT | CTAGAATCTA | 11220 |
| TAGGTAGTTT | GACACAAGAA | TTAGAATATA | GAGGTGAAAG | TCTATTATGC | AGTTTAATAT | 11280 |
| TTAGAAATGT | ATGGTTATAT | AATCAGATTG | CTCTACAATT | AAAAATCAT | GCATTATGTA | 11340 |
| ACAATAAACT | ATATTTGGAC | ATATTAAAGG | TTCTGAAACA | CTTAAAAACC | TTTTTTAATC | 11400 |
| TTGATAATAT | TGATACAGCA | TTAACATTGT | ATATGAATTT | ACCCATGTTA | TTTGGTGGTG | 11460 |
| GTGATCCCAA | CTTGTTATAT | CGAAGTTTCT | ATAGAAGAAC | TCCTGACTTC | CTCACAGAGG | 11520 |
| CTATAGTTCA | CTCTGTGTTT | ATACTTAGTT | ATTATACAAA | CCATGACTTA | AAAGATAAAC | 11580 |
| TTCAAGATCT | GTCAGATGAT | AGATTGAATA | AGTTCTTAAC | ATGCATAATC | ACGTTTGACA | 11640 |
| AAAACCCTAA | TGCTGAATTC | GTAACATTGA | TGAGAGATCC | TCAAGCTTTA | GGGTCTGAGA | 11700 |
| GACAAGCTAA | AATTACTAGC | GAAATCAATA | GACTGGCAGT | TACAGAGGTT | TTGAGTACAG | 11760 |
| CTCCAAACAA | AATATTCTCC | AAAAGTGCAC | AACATTATAC | TACTACAGAG | ATAGATCTAA | 11820 |
| ATGATATTAT | GCAAAATATA | GAACCTACAT | ATCCTCATGG | GCTAAGAGTT | GTTTATGAAA | 11880 |
| GTTTACCCTT | TTATAAAGCA | GAGAAAATAG | TAAATCTTAT | ATCAGGTACA | AAATCTATAA | 11940 |
| CTAACATACT | GGAAAAAACT | TCTGCCATAG | ACTTAACAGA | TATTGATAGA | GCCACTGAGA | 12000 |
| TGATGAGGAA | AAACATAACT | TTGCTTATAA | GGATACTTCC | ATTGGATTGT | AACAGAGATA | 12060 |
| AAAGAGAGAT | ATTGAGTATG | GAAAACCTAA | GTATTACTGA | ATTAAGCAAA | TATGTTAGGG | 12120 |
| AAAGATCTTG | GTCTTTATCC | AATATAGTTG | GTGTTACATC | ACCCAGTATC | ATGTATACAA | 12180 |
| TGGACATCAA | ATATACTACA | AGCACTATAT | CTAGTGGCAT | AATTATAGAG | AAATATAATG | 12240 |
| TTAACAGTTT | AACACGTGGT | GAGAGAGGAC | CCACTAAACC | ATGGGTGGT | TCATCTACAC | 12300 |
| AAGAGAAAAA | AACAATGCCA | GTTTATAATA | GACAAGTCTT | AACCAAAAAA | CAGAGAGATC | 12360 |
| AAATAGATCT | ATTAGCAAAA | TTGGATTGGG | TGTATGCATC | TATAGATAAC | AAGGATGAAT | 12420 |
| TCATGGAAGA | ACTCAGCATA | GGAACCCCTG | GGTTAACATA | TGAAAAGGCC | AAGAAATTAT | 12480 |
| TTCCACAATA | TTTAAGTGTC | AATTATTTGC | ATCGCCTTAC | AGTCAGTAGT | AGACCATGTG | 12540 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-------|
| AATTCCCTGC | ATCAATACCA | GCTTATAGAA | CAACAAATTA | TCACTTTGAC | ACTAGCCCTA | 12600 |
| TTAATCGCAT | ATTAACAGAA | AAGTATGGTG | ATGAAGATAT | TGACATAGTA | TTCCAAAAC | 12660 |
| GTATAAGCTT | TGGCCTTAGT | TTAATGTCAG | TAGTAGAACA | ATTTACTAAT | GTATGTCCTA | 12720 |
| ACAGAATTAT | TCTCATACCT | AAGCTTAATG | AGATACATTT | GATGAAACCT | CCCATATTCA | 12780 |
| CAGGTGATGT | TGATATTCAC | AAGTTAAAC | AAGTGATACA | AAAACAGCAT | ATGTTTTTAC | 12840 |
| CAGACAAAAT | AAGTTTGACT | CAATATGTGG | AATTATTCTT | AAGTAATAAA | AACTCAAAT | 12900 |
| CTGGATCTCA | TGTTAATTCT | AATTTAATAT | TGGCACATAA | AATATCTGAC | TATTTTCATA | 12960 |
| ATACTTACAT | TTTAAGTACT | AATTTAGCTG | GACATTGGAT | TCTGATTATA | CAACTTATGA | 13020 |
| AAGATTCTAA | AGGTATTTTT | GAAAAAGATT | GGGGAGAGGG | ATATATAACT | GATCATATGT | 13080 |
| TTATTAATTT | GAAAGTTTTT | TTCAATGCTT | ATAAGACCTA | TCTCTTGTTG | TTTCATAAAG | 13140 |
| GTTATGGCAA | AGCAAAGCTG | GAGTGTGATA | TGAACACTTC | AGATCTTCTA | TGTGTATTGG | 13200 |
| AATTAATAGA | CAGTAGTTAT | TGGAAGCTA | TGTCTAAGGT | ATTTTTAGAA | CAAAAAGTTA | 13260 |
| TCAAATACAT | TCTTAGCCAA | GATGCAAGTT | TACATAGAGT | AAAAGGATGT | CATAGCTTCA | 13320 |
| AATTATGGTT | TCTTAAACGT | CTTAATGTAG | CAGAATTCAC | AGTTTGCCCT | TGGGTTGTGA | 13380 |
| ACATAGATTA | TCATCCAACA | CATATGAAAG | CAATATTAAC | TTATATAGAT | CTTGTTAGAA | 13440 |
| TGGGATTGAT | AAATATAGAT | AGAATACACA | TTAAAAATAA | ACACAAATTC | AATGATGAAT | 13500 |
| TTTATACTTC | TAATCTCTTC | TACATTAATT | ATAACTTCTC | AGATAATACT | CATCTATTAA | 13560 |
| CTAAACATAT | AAGGATTGCT | AATCTGAAT | TAGAAAATAA | TTACAACAAA | TTATATCATC | 13620 |
| CTACACCAGA | AACCCTAGAG | AATATACTAG | CCAATCCGAT | TAAAAGTAAT | GACAAAAAGA | 13680 |
| CACTGAATGA | CTATTGTATA | GGTAAAAATG | TTGACTCAAT | AATGTTACCA | TTGTTATCTA | 13740 |
| ATAAGAAGCT | TATTAAATCG | TCTGCAATGA | TTAGAACCAA | TTACAGCAAA | CAAGATTTGT | 13800 |
| ATAATTTATT | CCCTATGGTT | GTGATTGATA | GAATTATAGA | TCATTCAGGC | AATACAGCCA | 13860 |
| AATCCAACCA | ACTTTTAACT | ACTACTTCCC | ACCAAATATC | CTTAGTGCAC | AATAGCACAT | 13920 |
| CACTTTACTG | CATGCTTCCT | TGGCATCATA | TTAATAGATT | CAATTTTGTA | TTTAGTTCTA | 13980 |
| CAGGTTGTAA | AATTAGTATA | GAGTATATTT | TAAAAGATCT | TAAAATTAAA | GATCCCAATT | 14040 |
| GTATAGCATT | CATAGGTGAA | GGAGCAGGGA | ATTTATTATT | GCGTACAGTA | GTGGAACCTC | 14100 |
| ATCCTGACAT | AAGATATATT | TACAGAAGTC | TGAAAGATTG | CAATGATCAT | AGTTTACCTA | 14160 |
| TTGAGTTTTT | AAGGCTGTAC | AATGGACATA | TCAACATTGA | TTATGGTGAA | AATTTGACCA | 14220 |
| TTCTGTCTAC | AGATGCAACC | AACAACATTC | ATTGGTCTTA | TTTACATATA | AAGTTTGCTG | 14280 |
| AACCTATCAG | TCTTTTTGTC | TGTGATGCCG | AATTGTCTGT | AACAGTCAAC | TGGAGTAAAA | 14340 |
| TTATAATAGA | ATGGAGCAAG | CATGTAAGAA | AGTGCAAGTA | CTGTTCTCTA | GTTAATAAAT | 14400 |
| GTATGTTAAT | AGTAAAATAT | CATGCTCAAG | ATGATATTGA | TTTCAAATTA | GACAATATAA | 14460 |
| CTATATTAAA | AACCTATGTA | TGCTTAGGCA | GTAAGTTAAA | GGGATCGGAG | GTTTACTTAG | 14520 |
| TCCTTACAAT | AGGTCCTGCG | AATATATTCC | CAGTATTTAA | TGTAGTACAA | AATGCTAAAT | 14580 |
| TGATACTATC | AAGAACCAAA | AATTCATCA | TGCCTAAGAA | AGCTGATAAA | GAGTCTATTG | 14640 |

| | | | | | | |
|------------|------------|-------------|------------|-------------|-------------|-------|
| ATGCAAATAT | TAAAAGTTTG | ATACCCTTTC | TTTGTTACCC | TATAACAAAA | AAAGGAATTA | 14700 |
| ATACTGCATT | GTCAAAACTA | AAGAGTGTG | TTAGTGGAGA | TATACTATCA | TATTCTATAG | 14760 |
| CTGGACGTAA | TGAAGTTTTC | AGCAATAAAC | TTATAAATCA | TAAGCATATG | AACATCTTAA | 14820 |
| AATGGTTCAA | TCATGTTTTA | AATTTTCAGAT | CAACAGAACT | AAACTATAAC | CATTTTATATA | 14880 |
| TGGTAGAATC | TACATATCCT | TACCTAAGTG | AATTGTTAAA | CAGCTTGACA | ACCAATGAAC | 14940 |
| TTAAAAAACT | GATTAAAATC | ACAGGTAGTC | TGTTATACAA | CTTTCATAAT | GAATAATGAA | 15000 |
| TAAAGATCTT | ATAATAAAAA | TTCCCATAGC | TATACACTAA | CACTGTATTTC | AATTATAGTT | 15060 |
| ATTAAAAATT | AAAAATCATA | TAATTTTTTTA | AATAACTTTT | AGTGAACATA | TCCTAAAGTT | 15120 |
| ATCATTTTAA | TCTTGGAGGA | ATAAATTTAA | ACCCTAATCT | AATTGGTTTA | TATGTGTATT | 15180 |
| AACTAAATTA | CGAGATATTA | GTTTTTGACA | CTTTTTTTCT | CGT | | 15223 |

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15225 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| ACGCGAAAAA | ATGCGTACTA | CAAACTTGCA | CATTCGGAAA | AAATGGGGCA | AATAAGAATT | 60 |
| TGATAAGTGC | TATTTAAGTC | TAACCTTTTC | AATCAGAAAT | GGGGTGCAAT | TCACTGAGCA | 120 |
| TGATAAAGGT | TAGATTACAA | AATTTATTTG | ACAATGACGA | AGTAGCATTG | TTAAAAATAA | 180 |
| CATGTTATAC | TGACAAATTA | ATTCTTCTGA | CCAATGCATT | AGCCAAAGCA | GCAATACATA | 240 |
| CAATTAAATT | AAACGGTATA | GTTTTTATAC | ATGTTATAAC | AAGCAGTGAA | GTGTGCCCTG | 300 |
| ATAACAACAT | TGTAGTAAAA | TCTAACTTTA | CAACAATGCC | AATATTACAA | AACGGAGGAT | 360 |
| ACATATGGGA | ATTGATTGAG | TTGACACACT | GCTCTCAATT | AAACGGTCTA | ATGGATGATA | 420 |
| ATTGTGAAAT | CAAATTTTCT | AAAAGACTAA | GTGACTCAGT | AATGACTAAT | TATATGAATC | 480 |
| AAATATCTGA | TTTACTTGGG | CTTGATCTCA | ATTCATGAAT | TATGTTTAGT | CTAACTCAAT | 540 |
| AGACATGTGT | TTATTACCAT | TTTAGTTAAT | ATAAAAACTC | ATCAAAGGGA | AATGGGGCAA | 600 |
| ATAAACTCAC | CTAATCAATC | AAACTATGAG | CACTACAAAT | GACAACACTA | CTATGCAAAG | 660 |
| ATTAATGATC | ACGGACATGA | GACCCCTGTC | GATGGATTCA | ATAATAACAT | CTCTCACCAA | 720 |
| AGAAATCATC | ACACACAAAT | TCATATACTT | GATAACAAT | GAATGTATTG | TAAGAAAACT | 780 |
| TGATGAAAGA | CAAGCTACAT | TTACATTCTT | AGTCAATTAT | GAGATGAAGC | TACTGCACAA | 840 |
| AGTAGGGAGT | ACCAAATACA | AGAAATACAC | TGAATATAAT | ACAAAATATG | GCACTTTCCC | 900 |
| CATGCCTATA | TTTATCAATC | ATGGCGGGTT | TCTAGAATGT | ATTGGCATT | AGCCTACAAA | 960 |
| ACACACTCCT | ATAATATACA | AATATGACCT | CAACCCGTAA | ATTCCAACAA | AAAAAACCAA | 1020 |
| CCCAACCAAA | CCAAGCTATT | CCTCAAACAA | CAATGCTCAA | TAGTTAAGAA | GGAGCTAATC | 1080 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| CGTTTTAGTA | ATTAAAAATA | AAAGTAAAGC | CAATAACATA | AATTGGGGCA | AATACAAAGA | 1140 |
| TGGCTCTTAG | CAAAGTCAAG | TTAAATGATA | CATTAAATAA | GGATCAGCTG | CTGTCATCCA | 1200 |
| GCAAATACAC | TATTCAACGT | AGTACAGGAG | ATAATATTGA | CACTCCCAAT | TATGATGTGC | 1260 |
| AAAAACACCT | AAACAAACTA | TGTGGTATGC | TATTAATCAC | TGAAGATGCA | AATCATAAAT | 1320 |
| TCACAGGATT | AATAGGTATG | TTATATGCTA | TGTCCAGGTT | AGGAAGGGAA | GACACTATAA | 1380 |
| AGATACTTAA | AGATGCTGGA | TATCATGTTA | AAGCTAATGG | AGTAGATATA | ACAACATATC | 1440 |
| GTCAAGATAT | AAATGGAAAG | GAAATGAAAT | TCGAAGTATT | AACATTATCA | AGCTTGACAT | 1500 |
| CAGAAATACA | AGTCAATATT | GAGATAGAAT | CTAGAAAATC | CTACAAAAAA | ATGCTAAAAG | 1560 |
| AGATGGGAGA | AGTGGCTCCA | GAATATAGGC | ATGATTCTCC | AGACTGTGGG | ATGATAATAC | 1620 |
| TGTGTATAGC | AGCACTTGTA | ATAACCAAAT | TAGCAGCAGG | AGACAGATCA | GGTCTTACAG | 1680 |
| CAGTAATTAG | GAGGGCAAAC | AATGTCTTAA | AAAATGAAAT | AAAACGCTAC | AAGGGTCTCA | 1740 |
| TACCAAAGGA | TATAGCTAAC | AGTTTTTATG | AAGTGTTTGA | AAAACACCCT | CATCTTATAG | 1800 |
| ATGTTTTTGT | GCACTTTGGC | ATTGCACAAT | CATCAACAAG | AGGGGGTAGT | AGAGTTGAAG | 1860 |
| GAATCTTTGC | AGGATTGTTT | ATGAATGCCT | ATGGTTCAGG | GCAAGTAATG | CTAAGATGGG | 1920 |
| GAGTTTTAGC | CAAATCTGTA | AAAAATATCA | TGCTAGGTCA | TGCTAGTGTG | CAGGCAGAAA | 1980 |
| TGGAGCAAGT | TGTGGAAGTC | TATGAGTATG | CACAGAAGTT | GGGAGGAGAA | GCTGGATTCT | 2040 |
| ACCATATATT | GAACAATCCA | AAAGCATCAT | TGCTGTCATT | AACTCAATTT | CCTAACTTCT | 2100 |
| CAAGTGTGGT | CCTAGGCAAT | GCAGCAGGTC | TAGGCATAAT | GGGAGAGTAT | AGAGGTACGC | 2160 |
| CAAGAAACCA | GGATCTTTAT | GATGCAGCCA | AAGCATATGC | AGAGCAACTC | AAAGAAAATG | 2220 |
| GAGTAATAAA | CTACAGTGTA | TTAGACTTAA | CAGCAGAAGA | ATTGGAAGCC | ATAAAGAATC | 2280 |
| AACTCAACCC | TAAAGAAGAT | GATGTAGAGC | TTTAAGTTAA | CAAAAAATAC | GGGGCAAATA | 2340 |
| AGTCAACATG | GAGAAGTTTG | CACCTGAATT | TCATGGAGAA | GATGCAAATA | ACAAAGCTAC | 2400 |
| CAAATTCCTA | GAATCAATAA | AGGGCAAGTT | CGCATCATCC | AAAGATCCTA | AGAAGAAAGA | 2460 |
| TAGCATAATA | TCTGTAACT | CAATAGATAT | AGAAGTAACC | AAAGAGAGCC | CGATAACATC | 2520 |
| TGGCACCAAC | ATCATCAATC | CAACAAGTGA | AGCCGACAGT | ACCCCAAGAA | CCAAAGCCAA | 2580 |
| CTACCCAAGA | AAACCCCTAG | TAAGCTTCAA | AGAAGATCTC | ACCCCAAGTG | ACAACCCTTT | 2640 |
| TTCTAAGTTG | TACAAAGAAA | CAATAGAAAC | ATTTGATAAC | AATGAAGAAG | AATCTAGCTA | 2700 |
| CTCATATGAA | GAGATAAATG | ATCAAACAAA | TGACAACATT | ACAGCAAGAC | TAGATAGAAT | 2760 |
| TGATGAAAAA | TTAAGTGAAA | TATTAGGAAT | GCTCCATACA | TTAGTAGTTG | CAAGTGCAGG | 2820 |
| ACCCACTTCA | GCTCGCGATG | GAATAAGAGA | TGCTATGGTT | GGTCTGAGAG | AAGAAATGAT | 2880 |
| AGAAAAAATA | AGAGCGGAAG | CATTAATGAC | CAATGATAGG | TTAGAGGCTA | TGGCAAGACT | 2940 |
| TAGGAATGAG | GAAAGCGAAA | AAATGGCAAA | AGACACCTCA | GATGAAGTGC | CTCTTAATCC | 3000 |
| AACTTCCAAA | AAATTGAGTG | ACTTGTTGGA | AGACAACGAT | AGTGACAATG | ATCTGTCACT | 3060 |
| TGATGATTTT | TGATCAGTGA | TCAACTCACT | CAGCAATCAA | CAACATCAAT | AAAACAGACA | 3120 |
| TCAATCCATT | GAATCAACTG | CCAGACCGAA | CAAACAAATG | TCCGTCAGCG | GAACCACCAA | 3180 |

| | | | | | | |
|-------------|-------------|------------|------------|-------------|------------|------|
| CCAATCAATC | AACCAACTGA | TCCATCAGCA | ACCTGACGAA | ATTAACAATA | TAGTAACAAA | 3240 |
| AAAAGAACAA | GATGGGGCAA | ATATGGAAAC | ATACGTGAAC | AAGCTTCACG | AAGGCTCCAC | 3300 |
| ATACACAGCA | GCTGTTTCAGT | ACAATGTTCT | AGAAAAAGAT | GATGATCCTG | CATCACTAAC | 3360 |
| AATATGGGTG | CCTATGTTCC | AGTCATCTGT | ACCAGCAGAC | TTGCTCATAA | AAGAACTTGC | 3420 |
| AAGCATCAAC | ATACTAGTGA | AGCAGATCTC | TACGCCCAAA | GGACCTTCAC | TACGAGTCAC | 3480 |
| GATTAACTCA | AGAAGTGCTG | TGCTGGCTCA | AATGCCTAGT | AATTTTCATCA | TAAGCGCAAA | 3540 |
| TGTATCATT | GATGAAAGAA | GCAAATTAGC | ATATGATGTA | ACTACACCTT | GTGAAATCAA | 3600 |
| AGCATGCAGT | CTAACATGCT | TAAAAGTGAA | AAGTATGTTA | ACTACAGTCA | AAGATCTTAC | 3660 |
| CATGAAGACA | TTCAACCCCA | CTCATGAGAT | CATTGCTCTA | TGTGAATTTG | AAAATATTAT | 3720 |
| GACATCAAAA | AGAGTAATAA | TACCAACCTA | TCTAAGACCA | ATTAGTGTCA | AAAACAAGGA | 3780 |
| TCTGAACTCA | CTAGAAAACA | TAGCAACCAC | CGAATTCAAA | AATGCTATCA | CCAATGCGAA | 3840 |
| AATTATTCCC | TATGCTGGAT | TAGTATTAGT | TATCACAGTT | ACTGACAATA | AAGGAGCATT | 3900 |
| CAAATATATC | AAGCCACAGA | GTCAATTTAT | AGTAGATCTT | GGTGCCTACC | TAGAAAAAGA | 3960 |
| GAGCATATAT | TATGTGACTA | CTAATTGGAA | GCATACAGCT | ACACGTTTTT | CAATCAAACC | 4020 |
| ACTAGAGGAT | TAAATTTAAT | TATCAACACT | GAATGACAGG | TCCACATATA | TCCTCAAAC | 4080 |
| ACACACTATA | TCCAAACATC | ATGAACATCT | ACACTACACA | CTTCATCACA | CAAACCAATC | 4140 |
| CCACTCAAAA | TCCAAATCA | CTACCAGCCA | CTATCTGCTA | GACCTAGAGT | GCGAATAGGT | 4200 |
| AAATAAAAACC | AAAATATGGG | GTAAATAGAC | ATTAGTTAGA | GTTCAATCAA | TCTCAACAAC | 4260 |
| CATTATATACC | GCCAATTCAA | TACATATACT | ATAAATCTTA | AAATGGGAAA | TACATCCATC | 4320 |
| ACAATAGAAT | TCACAAGCAA | ATTTTGCCCC | TATTTTACAC | TAATACATAT | GATCTTAACT | 4380 |
| CTAATCTCTT | TACTAATTAT | AATCACTATT | ATGATTGCAA | TACTAAATAA | GCTAAGTGAA | 4440 |
| CATAAAACAT | TCTGTAACAA | TACTCTTGAA | CTAGGACAGA | TGCATCAAAT | CAACACATAG | 4500 |
| TGCTCTACCA | TCATGCTGTG | TCAAATTATA | ATCCTGTATA | TATAAACAAA | CAAATCCAAT | 4560 |
| CTTCTCACAG | AGTCATGGTG | TCGCAAAACC | ACGCCAACTA | TCATGGTAGC | ATAGAGTAGT | 4620 |
| TATTTAAAAA | TTAACATAAT | GATGAATTAT | TAGTATGGGA | TCAAAAACAA | CATTGGGGCA | 4680 |
| AATGCAACCA | TGTCCAAACA | CAAGAATCAA | CGCACTGCCA | GGACTCTAGA | AAAGACCTGG | 4740 |
| GATACTCTCA | ATCATCTAAT | TGTAATATCC | TCTTGTTTAT | ACAGATTAAA | TTTAAATCT | 4800 |
| ATAGCACAAA | TAGCACTATC | AGTTCTGGCA | ATGATAATCT | CAACCTCTCT | CATAATTGCA | 4860 |
| GCCATAATAT | TCATCATCTC | TGCCAATCAC | AAAGTTACAC | TAACAACGGT | CACAGTTCAA | 4920 |
| ACAATAAAAA | ACCACACTGA | AAAAAACATC | ACCACCTACC | TTACTCAAGT | CCCACCAGAA | 4980 |
| AGGGTTAGCT | CATCCAAACA | ACCTACAACC | ACATCACCAA | TCCACACAAA | TTCAGCCACA | 5040 |
| ACATCACCCA | ACACAAAGTC | AGAAACACAC | CACACAACAG | CACAAACCAA | AGGCAGAACC | 5100 |
| ACCACCTCAA | CACAGACCAA | CAAGCCGAGC | ACAAAACCAC | GCCTAAAAAA | TCCACCAAAA | 5160 |
| AAACCAAAAG | ATGATTACCA | TTTTGAAGTG | TTCAACTTCG | TTCCCTGTAG | TATATGTGGC | 5220 |
| AACAATCAAC | TTTGCAAATC | CATCTGTAAA | ACAATACCAA | GCAACAAACC | AAAGAAGAAA | 5280 |

| | | | | | | |
|------------|------------|-------------|-------------|-------------|------------|------|
| CCAACCATCA | AACCCACAAA | CAAACCAACC | ACCAAAACCA | CAAACAAAAG | AGACCCAAAA | 5340 |
| ACACCAGCCA | AAACGACGAA | AAAAGAAACT | ACCACCAACC | CAACAAAAAA | ACCAACCCTC | 5400 |
| ACGACCACAG | AAAGAGACAC | CAGCACCTCA | CAATCCACTG | TGCTCGACAC | AACCACATTA | 5460 |
| GAACACACAA | TCCAACAGCA | ATCCCTCCAC | TCAACCACCC | CCGAAAACAC | ACCCAACTCC | 5520 |
| ACACAAACAC | CCACAGCATC | CGAGCCCTCT | ACATCAAATT | CCACCCAAAA | TACCCAATCA | 5580 |
| CATGCTTAGT | TATTCAAAAA | CTACATCTTA | GCAGAAAACC | GTGACCTATC | AAGCAAGAAC | 5640 |
| GAAATTAAAC | CTGGGGCAAA | TAACCATGGA | GCTGCTGATC | CACAGGTAA | GTGCAATCTT | 5700 |
| CCTAACTCTT | GCTATTAATG | CATTGTACCT | CACCTCAAGT | CAGAACATAA | CTGAGGAGTT | 5760 |
| TTACCAATCG | ACATGTAGTG | CAGTTAGCAG | AGGTTATTTT | AGTGCTTTAA | GAACAGGTTG | 5820 |
| GTATACCAGT | GTCATAACAA | TAGAATTAAG | TAATATAAAA | GAAACCAAAT | GCAATGGAAC | 5880 |
| TGACACTAAA | GTAAAACTTA | TAAAACAAGA | ATTAGATAAG | TATAAGAATG | CAGTGACAGA | 5940 |
| ATTACAGCTA | CTTATGCAAA | ACACACCAGC | TGCCAACAAAC | CGGGCCAGAA | GAGAAGCACC | 6000 |
| ACAGTATATG | AACTATACAA | TCAATACCAC | TAAAAACCTA | AATGTATCAA | TAAGCAAGAA | 6060 |
| GAGGAAACGA | AGATTTCTGG | GCTTCTTGTT | AGGTGTAGGA | TCTGCAATAG | CAAGTGGTAT | 6120 |
| AGCTGTATCC | AAAGTTCTAC | ACCTTGAAGG | AGAAGTGAAC | AAGATCAAAA | ATGCTTTGTT | 6180 |
| ATCTACAAAC | AAAGCTGTAG | TCAGTCTATC | AAATGGGGTC | AGTGTTTTAA | CCAGCAAAGT | 6240 |
| GTTAGATCTC | AAGAATTACA | TAAATAACCA | ATTATTACCC | ATAGTAAATC | AACAGAGCTG | 6300 |
| TCGCATCTCC | AACATTGAAA | CAGTTATAGA | ATTCCAGCAG | AAGAACAGCA | GATTGTTGGA | 6360 |
| AATCAACAGA | GAATTCACTG | TCAATGCAGG | TGTAACAACA | CCTTTAAGCA | CTTACATGTT | 6420 |
| AACAAACAGT | GAGTTACTAT | CATTGATCAA | TGATATGCCT | ATAACAAATG | ATCAGAAAAA | 6480 |
| ATTAATGTCA | AGCAATGTTT | AGATAGTAAG | GCAACAAAGT | TATTCTATCA | TGTCTATAAT | 6540 |
| AAAGGAAGAA | GTCCTTGCAT | ATGTTGTACA | GCTACCTATC | TATGGTGTAA | TAGATACACC | 6600 |
| TTGCTGGAAA | TTACACACAT | CACCTCTATG | CACCACCAAC | ATCAAAGAAG | GATCAAATAT | 6660 |
| TTGTTTAAAC | AGGACTGATA | GAGGATGGTA | TTGTGATAAT | GCAGGATCAG | TATCCTTCTT | 6720 |
| TCCACAGGCT | GACACTTGTA | AAGTACAGTC | CAATCGAGTA | TTTTGTGACA | CTATGAACAG | 6780 |
| TTTGACATTA | CCAAGTGAAG | TCAGCCTTTG | TAACACTGAC | ATATTCAATT | CCAAGTATGA | 6840 |
| CTGCAAAATT | ATGACATCAA | AAACAGACAT | AAGCAGCTCA | GTAATTACTT | CTCTTGAGAC | 6900 |
| TATAGTGTCA | TGCTATGGTA | AACTAAATG | CACTGCATCC | AACAAAAATC | GTGGGATTAT | 6960 |
| AAAGACATTT | TCTAATGGTT | GTGACTATGT | GTCAAACAAA | GGAGTAGATA | CTGTGTCAGT | 7020 |
| GGGCAACACT | TTATACTATG | TAAACAAGCT | GGAAGGCAAG | AACCTTTATG | TAAAAGGGGA | 7080 |
| ACCTATAATA | AATTACTATG | ACCCTCTAGT | GTTTCCTTCT | GATGAGTTTG | ATGCATCAAT | 7140 |
| ATCTCAAGTC | AATGAAAAAA | TCAATCAAAG | TTAGCTTTT | ATTCTGTAGAT | CTGATGAATT | 7200 |
| ACTACATAAT | GTAAATACTG | GCAAATCTAC | TACAAATATT | ATGATAACTA | CAATTATTAT | 7260 |
| AGTAATCATT | GTAGTATTGT | TATCATTAAAT | AGCTATTGGT | TTGCTGTTGT | ATTGCAAAGC | 7320 |
| CAAAAACACA | CCAGTTACAC | TAAGCAAAGA | CCAACTAAGT | GGAATCAATA | ATATTGCATT | 7380 |

| | | | | | | |
|------------|-------------|------------|------------|------------|-------------|------|
| CAGCAAATAG | ACAAAAAACC | ACCTGATCAT | GTTTCAACAA | CAGTCTGCTG | ATCACCAATC | 7440 |
| CCAAATCAAC | CCATAACAAA | CACTTCAACA | TCACAGTACA | GGCTGAATCA | TTTCTTCACA | 7500 |
| TCATGCTACC | CACACAAC TA | AGCTAGATCC | TTAACTCATA | GTTACATAAA | AACCTCAAGT | 7560 |
| ATCACAATCA | AACACTAAAT | CAACACATCA | TTCACAAAAT | TAACAGCTGG | GGCAAATATG | 7620 |
| TCGCGAAGAA | ATCCTTGTA | ATTTGAGATT | AGAGGTCATT | GCTTGAATGG | TAGAAGATGT | 7680 |
| CACTACAGTC | ATAATTACTT | TGAATGGCCT | CCTCATGCCT | TACTAGTGAG | GCAAACTTC | 7740 |
| ATGTTAAACA | AGATACTCAA | GTCAATGGAC | AAAAGCATAG | ACACTTTGTC | TGAAATAAGT | 7800 |
| GGAGCTGCTG | AACTGGACAG | AACAGAAGAA | TATGCTCTTG | GTATAGTTGG | AGTGCTAGAG | 7860 |
| AGTTACATAG | GATCTATAAA | CAACATAACA | AAACAATCAG | CATGTGTTGC | TATGAGTAAA | 7920 |
| CTTCTTATTG | AGATCAATAG | TGATGACATT | AAAAAGCTGA | GAGATAATGA | AGAACCCAAT | 7980 |
| TCACCTAAGA | TAAGAGTGTA | CAATACTGTT | ATATCATACA | TTGAGAGCAA | TAGAAAAAAC | 8040 |
| AACAAGCAAA | CAATCCATCT | GCTCAAAAGA | CTACCAGCAG | ACGTGCTGAA | GAAGACAATA | 8100 |
| AAAAACACAT | TAGATATCCA | CAAAAGCATA | ATCATAAGCA | ACCCAAAAGA | GTCAACCGTG | 8160 |
| AATGATCAAA | ATGACCAAAC | CAAAAATAAT | GATATTACCG | GATAAATATC | CTTG TAGTAT | 8220 |
| ATCATCCATA | TTGATTTCAA | GTGAAAGCAT | GATTGCTACA | TTCAATCATA | AAAACATATT | 8280 |
| ACAATTTAAC | CATAACCATT | TGGATAACCA | CCAGCGTTTA | TTAAATAATA | TATTTGATGA | 8340 |
| AATTCATTGG | ACACCTAAAA | ACTTATTAGA | TGCCACTCAA | CAATTTCTCC | AACATCTTAA | 8400 |
| CATCCCTGAA | GATATATATA | CAATATATAT | ATTAGTGTC | TAATGCTTGG | CCATAACGAT | 8460 |
| TCTATATCAT | CCAACCATAA | AACTATCTTA | ATAAGGTTAT | GGGACAAAAT | GGATCCCATT | 8520 |
| ATTAATGGAA | ACTCTGCTAA | TGTGTATCTA | ACTGATAGTT | ATTTAAAAGG | TGTTATCTCT | 8580 |
| TTTTCAGAAT | GTAATGCTTT | AGGGAGTTAC | CTTTTAAACG | GCCCTTATCT | CAAAAATGAT | 8640 |
| TACACCAACT | TAATTAGTAG | ACAAAGTCCA | CTACTAGAGC | ATATGAATCT | TAAAAAACTA | 8700 |
| ACTATAACAC | AGTCATTAA | ATCTAGATAT | CATAAAGGTG | AACTGAAATT | AGAAGAACCA | 8760 |
| ACTTATTTCC | AGTCATTACT | TATGACATAT | AAAAGCATGT | CCTCGTCTGA | ACAAATTGCT | 8820 |
| ACAACTAACT | TACTTAAAAA | AATAATACGA | AGAGCTATAG | AAATAAGTGA | TGTAAAGGTG | 8880 |
| TACGCCATCT | TGAATAAACT | AGGACTAAAG | GAAAAGGACA | GAGTTAAGCC | CAACAATAAT | 8940 |
| TCAGGTGATG | AAAACCTCAGT | ACTTACAAC | ATAATTAAAG | ATGATATACT | TTCCGCTGTG | 9000 |
| GAAAGCAATC | AATCATATAC | AAATTCAGAC | AAAAATCACT | CAGTAAATCA | AAATATCACT | 9060 |
| ATCAAAACAA | CACTCTTGAA | AAAATTGATG | TGTTCAATGC | AACATCCTCC | ATCATGGTTA | 9120 |
| ATACACTGGT | TCAATTTATA | TACAAAATTA | AATAACATAT | TAACACAATA | TCGATCAAAT | 9180 |
| GAGGTAAAAA | GTCATGGGTT | TATATTAATA | GATAATCAAA | CTTTAAGTGG | TTTTCAGTTT | 9240 |
| ATTTTAAATC | AATATGGTTG | TATCGTTTAT | CATAAAGGAC | TCAAAAAAAT | CACAACTACT | 9300 |
| ACTTACAATC | AATTTTAAAC | ATGGAAAGAC | ATCAGCCTTA | GCAGATTAAA | TGTTTGCTTA | 9360 |
| ATTACTTGGA | TAAGTAATTG | TTTGAATACA | TTAAATAAAA | GCTTAGGGCT | GAGATGTGGA | 9420 |
| TTCAATAATG | TTGTGTTATC | ACAATTATTT | CTTTATGGAG | ATTGTATACT | GAAATTATTT | 9480 |

| | | | | | | |
|------------|------------|------------|------------|-------------|-------------|-------|
| CATAATGAAG | GCTTCTACAT | AATAAAAGAA | GTAGAGGGAT | TTATTATGTC | TTTAATTCTA | 9540 |
| AACATAACAG | AAGAAGATCA | ATTTAGGAAA | CGATTTTATA | ATAGCATGCT | AAATAACATC | 9600 |
| ACAGATGCAG | CTATTAAGGC | TCAAAAGAAC | CTACTATCAA | GGGTATGTCA | CACTTTATTA | 9660 |
| GACAAGACAG | TGTCTGATAA | TATCATAAAT | GGTAAATGGA | TAATCCTATT | AAGTAAATTT | 9720 |
| CTTAAATTGA | TTAAGCTTGC | AGGTGATAAT | AATCTCAATA | ATTTGAGTGA | GCTATATTTT | 9780 |
| CTCTTCAGAA | TCTTTGGACA | TCCAATGGTT | GATGAAAGAC | AAGCAATGGA | TGCTGTAAGA | 9840 |
| ATTAACTGTA | ATGAAACTAA | GTTCTACTTA | TTAAGTAGTC | TAAGTACGTT | AAGAGGTGCT | 9900 |
| TTCATTTATA | GAATCATAAA | AGGGTTTGTA | AATACCTACA | ACAGATGGCC | CACTTTAAGG | 9960 |
| AATGCTATTG | TCCTACCTCT | AAGATGGTTA | AACTATTATA | AACTTAATAC | TTATCCATCT | 10020 |
| CTACTTGAAA | TCACAGAAAA | TGATTTGATT | ATTTTATCAG | GATTGCGGTT | CTATCGTGAA | 10080 |
| TTTCATCTGC | CTAAAAAAGT | GGATCTTGAA | ATGATAATAA | ATGACAAAGC | CATTTACCTT | 10140 |
| CCAAAAGATC | TAATATGGAC | TAGTTTTTCT | AGAAATTACA | TGCCATCACA | TATACAAAAT | 10200 |
| TATATAGAAC | ATGAAAAGTT | GAAGTTCTCT | GAAAGCGACA | GATCAAGAAG | AGTACTAGAG | 10260 |
| TATTACTTGA | GAGATAATAA | ATTCAATGAA | TGCGATCTAT | ACAATTGTGT | AGTCAATCAA | 10320 |
| AGCTATCTCA | ACAACTCTAA | TCACGTGGTA | TCACTAACTG | GTAAAGAAAG | AGAGCTCAGT | 10380 |
| GTAGGTAGAA | TGTTTGCTAT | GCAACCAGGT | ATGTTTAGGC | AAATCCAAAT | CTTAGCAGAG | 10440 |
| AAAATGATAG | CCGAAAATAT | TTTACAATTC | TTCCCTGAGA | GTTTGACAAG | ATATGGTGAT | 10500 |
| CTAGAGCTTC | AAAAGATATT | AGAATTAAAA | GCAGGAATAA | GCAACAAGTC | AAATCGTTAT | 10560 |
| AATGATAACT | ACAACAATTA | TATCAGTAAA | TGTTCTATCA | TTACAGATCT | TAGCAAATTC | 10620 |
| AATCAAGCAT | TTAGATATGA | AACATCATGT | ATCTGCAGTG | ATGTATTAGA | TGAAGTGCAT | 10680 |
| GGAGTACAAT | CTCTGTTCTC | TTGGTTGCAT | TTAACAATAC | CTCTTGTCAC | AATAATATGT | 10740 |
| ACATATAGAC | ATGCACCTCC | TTTCATAAAG | GATCATGTTG | TTAATCTTAA | TGAAGTTGAT | 10800 |
| GAACAAAGTG | GATTATACAG | ATATCATATG | GGTGGTATTG | AGGGCTGGTG | TCAAAAAC TG | 10860 |
| TGGACCATTG | AAGCTATATC | ATTATTAGAT | CTAATATCTC | TCAAAGGGAA | ATTCTCTATC | 10920 |
| ACAGCTCTGA | TAAATGGTGA | TAATCAGTCA | ATTGATATAA | GTAAACCAAGT | TAGACTTATA | 10980 |
| GAGGGTCAGA | CCCATGCTCA | AGCAGATTAT | TTGTTAGCAT | TAAATAGCCT | TAAATTGCTA | 11040 |
| TATAAAGAGT | ATGCAGGTAT | AGGCCATAAG | CTTAAGGGAA | CAGAGACCTA | TATATCCCGA | 11100 |
| GATATGCAGT | TCATGAGCAA | AACAATCCAG | CACAATGGAG | TGTACTATCC | AGCCAGTATC | 11160 |
| AAAAAAGTCC | TGAGAGTAGG | TCCATGGATA | AATACAATAC | TTGATGATTT | TAAAGTTAGT | 11220 |
| TTAGAATCTA | TAGGTAGCTT | AACACAGGAG | TTAGAATACA | GAGGGGAAAG | CTTATTATGC | 11280 |
| AGTTTAATAT | TTAGGAACAT | TTGGTTATAC | AATCAAATTG | CTTTGCAACT | CCGAAATCAT | 11340 |
| GCATTATGTA | ACAATAAGCT | ATATTTAGAT | ATATTGAAAG | TATTAACAACA | CTTAAAAACT | 11400 |
| TTTTTTAATC | TTGATAGTAT | CGATATGGCG | TTATCATTGT | ATATGAATTT | GCCTATGCTG | 11460 |
| TTTGGTGGTG | GTGATCCTAA | TTTGTTATAT | CGAAGCTTTT | ATAGGAGAAC | TCCAGACTTC | 11520 |
| CTTACAGAAG | CTATAGTACA | TTCAAGTGT | GTGTTGAGCT | ATTATACTGG | TCACGATTTA | 11580 |

| | | | | | | |
|-------------|------------|------------|------------|------------|-------------|-------|
| CAAGATAAGC | TCCAGGATCT | TCCAGATGAT | AGACTGAACA | AATTCTTGAC | ATGTGTCATC | 11640 |
| ACATTCGATA | AAAATCCCAA | TGCCGAGTTT | GTAACATTGA | TGAGGGATCC | ACAGGCGTTA | 11700 |
| GGGTCTGAAA | GGCAAGCTAA | AATTACTAGT | GAGATTAATA | GATTAGCAGT | AACAGAAGTC | 11760 |
| TTAAGTATAG | CTCCAAACAA | AATATTTTCT | AAAAGTGCAC | AACATTATAC | TACCACTGAG | 11820 |
| ATTGATCTAA | ATGACATTAT | GCAAAATATA | GAACCAACTT | ACCCTCATGG | ATTAAGAGTT | 11880 |
| GTTTATGAAA | GTCTACCTTT | TTATAAAGCA | GAAAAAATAG | TTAATCTTAT | ATCAGGAACA | 11940 |
| AAATCCATAA | CTAATATACT | TGAAAAAACA | TCAGCAATAG | ATACAACTGA | TATTAATAGG | 12000 |
| GCTACTGATA | TGATGAGGAA | AAATATAACT | TTACTTATAA | GGATACTTCC | ACTAGATTGT | 12060 |
| AACAAAGACA | AAAGAGAGTT | ATTAAGTTTA | GAAAATCTTA | GTATAACTGA | ATTAAGCAAG | 12120 |
| TATGTAAGAG | AAAGATCTTG | GTCATTATCC | AATATAGTAG | GAGTAACATC | GCCAAGTATT | 12180 |
| ATGTTACAA | TGGACATTAA | ATATACAACT | AGCACTATAG | CCAGTGGTAT | AATTATAGAA | 12240 |
| AAATATAATG | TTAATAGTTT | AATCGTGGT | GAAAGAGGAC | CTACTAAGCC | ATGGGTAGGT | 12300 |
| TCATCTACGC | AGGAGAAAAA | AACAATGCCA | GTGTACAATA | GACAAGTTTT | AACCAAAAAG | 12360 |
| CAAAGAGACC | AAATAGATTT | ATTAGCAAAA | TTAGACTGGG | TATATGCATC | CATAGACAAC | 12420 |
| AAAGATGAAT | TCATGGAAGA | ACTGAGTACT | GGAACACTTG | GACTGTCATA | TGAAAAAGCC | 12480 |
| AAAAAGTTGT | TTCCACAATA | TCTAAGTGTC | AATTATTTAC | ACCGTTTAAC | AGTCAGTAGT | 12540 |
| AGACCATGTG | AATTCCTGTC | ATCAATACCA | GCTTATAGAA | CAACAAATTA | TCATTTTCGAT | 12600 |
| ACTAGTCCTA | TCAATCATGT | ATTAACAGAA | AAGTATGGAG | ATGAAGATAT | CGACATTGTG | 12660 |
| TTTCAAAATT | GCATAAGTTT | TGGTCTTAGC | CTGATGTCGG | TTGTGGAACA | ATTCACAAAC | 12720 |
| ATATGTCCTA | ATAGAATTAT | TCTCATACCG | AAGCTGAATG | AGATACATTT | GATGAAACCT | 12780 |
| CCTATATTTA | CAGGAGATGT | TGATATCATC | AAGTTGAAGC | AAGTGATACA | AAAACAGCAT | 12840 |
| ATGTTCTAC | CAGATAAAAT | AAGTTTAACC | CAATATGTAG | AATTATTCCT | AAGTAACAAA | 12900 |
| GCACTTAAAT | CTGGATCTAA | CATCAATTCT | AATTTAATAT | TAGTACATAA | AATGTCTGAT | 12960 |
| TATTTTCATA | ATGCTTATAT | TTTAAGTACT | AATTTAGCTG | GACATTGGAT | TCTAATTATT | 13020 |
| CAACTTATGA | AAGATTCAAA | AGGTATTTTT | GAAAAAGATT | GGGGAGAGGG | GTACATAACT | 13080 |
| GATCATATGT | TCATTAATTT | GAATGTTTTT | TTAATGCTT | ATAAGACTTA | TTTGCTATGT | 13140 |
| TTTCATAAAG | GTTATGGTAA | AGCAAAATTA | GAATGTGATA | TGAACACTTC | AGATCTTCTT | 13200 |
| TGTGTTTTGG | AGTTAATAGA | CAGTAGCTAC | TGGAAATCTA | TGTCTAAAGT | TTTCCTAGAA | 13260 |
| CAAAAAGTCA | TAAAATACAT | AGTCAATCAA | GACACAAGTT | TGCATAGAAT | AAAAGGCTGT | 13320 |
| CACAGTTTTA | AGTTGTGGTT | TTTAAACGC | CTTAATAATG | CTAAATTTAC | CGTATGCCCT | 13380 |
| TGGGTGTGTTA | ACATAGATTA | TCACCCAACA | CATATGAAAG | CTATATTATC | TTACATAGAT | 13440 |
| TTAGTTAGAA | TGGGGTTAAT | AAATGTAGAT | AAATTAACCA | TTAAAAATAA | AAACAAATTC | 13500 |
| AATGATGAAT | TTTACACATC | AAATCTCTTT | TACATTAGTT | ATAACTTTTC | AGACAACACT | 13560 |
| CATTTGCTAA | CAAAACAAAT | AAGAATTGCT | AATTCAGAAT | TAGAAGATAA | TTATAACAAA | 13620 |
| CTATATCACC | CAACCCGAGA | AACTTTAGAA | AATATATCAT | TAATTCCTGT | TAAAAGTAAT | 13680 |

AATAGTAACA AACCTAAATT TTGTATAAGT GGAAATACCG AATCTATAAT GATGTCAACA 13740
 TTCTCTAATA AAATGCATAT TAAATCTTCC ACTGTTACCA CAAGATTCAA TTATAGCAAA 13800
 CAAGACTTGT ACAATTTATT TCCAAATGTT GTGATAGACA GGATTATAGA TCATTTCAGGT 13860
 AATACAGCAA AATCTAACCA ACTTTACATC ACCACTTCAC ATCAGACATC TTTAGTAAGG 13920
 AATAGTGCAT CACTTTATTG CATGCTTCCT TGGCATCATG TCAATAGATT TAACTTTGTA 13980
 TTTAGTTCCA CAGGATGCAA GATCAGTATA GAGTATATTT TAAAAGATCT TAAGATTAAAG 14040
 GACCCAGTT GTATAGCATT CATAGGTGAA GGAGCTGGTA ACTTATTATT ACGTACGGTA 14100
 GTAGAAGTTC ATCCAGACAT AAGATACATT TACAGAAGTT TAAAAGATTG CAATGATCAT 14160
 AGTTTACCTA TTGAATTTCT AAGATTATAC AACGGGCATA TAAACATAGA TTATGGTGAG 14220
 AATTTAACCA TTCCTGCTAC AGATGCAACT AATAACATTC ATTGGTCTTA TTTACATATA 14280
 AAATTTGCAG AACCTATTAG CATCTTTGTC TGCATGCTG AATTACCTGT TACAGCCAAT 14340
 TGGAGTAAAA TTATAATTGA ATGGAGTAAG CATGTAAGAA AGTGCAAGTA CTGTTCTTCT 14400
 GTAAATAGAT GCATTTTAAT CGCAAATAT CATGCTCAAG ATGATATTGA TTTCAAATTA 14460
 GATAACATTA CTATATTAAA AACTTACGTG TGCCTAGGTA GCAAGTTAAA AGGATCTGAA 14520
 GTTTACTTAG TCCTTACAAT AGGCCCTGCA AATATACTTC CTGTTTTTGA TGTTGTGCAA 14580
 AATGCTAAAT TGATTTTTTC AAGAACTAAA AATTTTATTA TGCCTAAAAA AACTGACAAG 14640
 GAATCTATCG ATGCAAATAT TAAAAGCTTA ATACCTTTCC TTTGTTACCC TATAACAAAA 14700
 AAAGGAATTA AGACTTCATT GTCAAAATTG AAGAGTGTAG TTAATGGGGA TATATTATCA 14760
 TATTCTATAG CTGGACGTAA TGAAGTATTC AGCAACAAGC TTATAAACCA CAAGCATATG 14820
 AATATCCTAA AATGGCTAGA TCATGTTTTA AATTTTAGAT CAGCTGAACT TAATTACAAT 14880
 CATTTATACA TGATAGAGTC CACATATCCT TACTTAAGTG AATTGTTAAA TAGTTTAAACA 14940
 ACCAATGAGC TCAAGAAACT GATTAAAATA ACAGGTAGTG TACTATACAA CCTTCCCAAC 15000
 GAACAGTAAC TTAAAATATC ATTAACAAGT TTGGTCAAAT TTAGATGCTA ACACATCATT 15060
 ATATTATAGT TATTAAAAAA TATGCAAACT TTTCAATAAT TTAGCTTACT GATTCCAAAA 15120
 TTATCATTTT ATTTTAAAGG GGTGAATAA AAGTCTAAAA CTAACAATGA TACATGTGCA 15180
 TTTACAACAC AACGAGACAT TAGTTTTTGA CACTTTTTTT CTCGT 15225

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACTCAAATAA GTTAATAAAA AATATCCCGG GAT

33

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCCGGGATAT TTTTATTAA CTTATTGAG T

31

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAAAGTATAT ATTATGTT

18

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TATATAAGCA CGATGATATG

20

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ACTCAAATAA GTTAAT

16

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TAACTTATTT GAGT

14

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GACACAACCC ACAATGATAA TACACCAC

28

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CATCTCTAAC CAAGGGAGTT AAATTTAAGT GG

32

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTAAGGAGAG ATATAAGATA GAAGATG

27

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTGTTTATATT AACTAATGGT GTTAGTG

27

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTATAATTGC AGCCATCATA TTCATAGCCT CGG

33

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GTGAAGTTGA GATTACAATT GCCAGAATGG

30